0570
1108

#4

OIPE

RAW SEQUENCE LISTING DATE: 11/21/2001
PATENT APPLICATION: US/09/818,066 TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt
Output Set: N:\CRF3\11212001\I818066.raw

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Shuping Tong et al.

(ii) TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR

(iii) NUMBER OF SEQUENCES: 75

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Fish & Richardson P.C.

(B) STREET: 225 Franklin Street

(C) CITY: Boston

(D) STATE: Massachusetts

(E) COUNTRY: U.S.A.

(F) ZIP: 02110-2804

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

C--> **(A) APPLICATION NUMBER: US/09/818,066**

C--> **(B) FILING DATE: 27-Mar-2001**

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/683,262

(B) FILING DATE: 18-JUL-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Fraser, Janis K.

(B) REGISTRATION NUMBER: 31,819

(C) REFERENCE/DOCKET NUMBER: 00786/287002

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 542-5070

(B) TELEFAX: (617) 542-8906

(C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCAGATCTAT GGGCAGAATC TTTCCAC 27

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

RAW SEQUENCE LISTING DATE: 11/21/2001
PATENT APPLICATION: US/09/818,066 TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt
Output Set: N:\CRF3\11212001\I818066.raw

```
68 GTGAATTCAG CGCAGGGTCC CCAAT                                          25
71 (2) INFORMATION FOR SEQ ID NO: 3:
73      (i) SEQUENCE CHARACTERISTICS:
74           (A) LENGTH: 28 base pairs
75           (B) TYPE: nucleic acid
76           (C) STRANDEDNESS: single
77           (D) TOPOLOGY: linear
79     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
81 TCAGATCTAT GATGGGGCAA CATCCAGC                                        28
83 (2) INFORMATION FOR SEQ ID NO: 4:
85      (i) SEQUENCE CHARACTERISTICS:
86           (A) LENGTH: 30 base pairs
87           (B) TYPE: nucleic acid
88           (C) STRANDEDNESS: single
89           (D) TOPOLOGY: linear
91     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
93 GCGAATTCAG GTACCAGACA TTTTCTTCTT                                      30
95 (2) INFORMATION FOR SEQ ID NO: 5:
97      (i) SEQUENCE CHARACTERISTICS:
98           (A) LENGTH: 27 base pairs
99           (B) TYPE: nucleic acid
100          (C) STRANDEDNESS: single
101          (D) TOPOLOGY: linear
103     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
105 GCGAATTCTT ATTCCTAACT CTTGTAA                                        27
107 (2) INFORMATION FOR SEQ ID NO: 6:
109      (i) SEQUENCE CHARACTERISTICS:
110          (A) LENGTH: 20 base pairs
111          (B) TYPE: nucleic acid
112          (C) STRANDEDNESS: single
113          (D) TOPOLOGY: linear
115     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
117 GARYTNTAYG TNATGGAGAT                                                20
119 (2) INFORMATION FOR SEQ ID NO: 7:
121      (i) SEQUENCE CHARACTERISTICS:
122          (A) LENGTH: 23 base pairs
123          (B) TYPE: nucleic acid
124          (C) STRANDEDNESS: single
125          (D) TOPOLOGY: linear
127     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
129 AAYTCNGGYT CNCCNGCYTC RTG                                            23
131 (2) INFORMATION FOR SEQ ID NO: 8:
133      (i) SEQUENCE CHARACTERISTICS:
134          (A) LENGTH: 21 base pairs
135          (B) TYPE: nucleic acid
136          (C) STRANDEDNESS: single
137          (D) TOPOLOGY: linear
139     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
141 TKYTNAGYCA YGARTTYCAR G                                              21
```

RAW SEQUENCE LISTING DATE: 11/21/2001
PATENT APPLICATION: US/09/818,066 TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt
Output Set: N:\CRF3\11212001\I818066.raw

(2) INFORMATION FOR SEQ ID NO: 9:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
TTKGCNGART ANARNGTYTC 20
(2) INFORMATION FOR SEQ ID NO: 10:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
ATGAAACAGA CACTGAAGAA 20
(2) INFORMATION FOR SEQ ID NO: 11:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
ATGGAGATCT CGGACGGCCC 20
(2) INFORMATION FOR SEQ ID NO: 12:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
TTCTTCAGTG TCTGTTTCAT 20
(2) INFORMATION FOR SEQ ID NO: 13:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
Gly Ser Arg Arg Ala Ser Val Gly Ser
1 5
(2) INFORMATION FOR SEQ ID NO: 14:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
ATCACTGAGC TCAAATTACC CCATGAGATG 30
(2) INFORMATION FOR SEQ ID NO: 15:

RAW SEQUENCE LISTING DATE: 11/21/2001
PATENT APPLICATION: US/09/818,066 TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt
Output Set: N:\CRF3\11212001\I818066.raw

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GGAAACTCGA GCTGGAAGCA GTGTTATGAA 30

(2) INFORMATION FOR SEQ ID NO: 16:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

ATGGTACCAT GGAGGCGGCG CGGTGCATCG AGC 33

(2) INFORMATION FOR SEQ ID NO: 17:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ATCTCGAGAT ATTAACATTA GCAATGTTAC T 31

(2) INFORMATION FOR SEQ ID NO: 18:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Gln Trp Thr Pro Glu Glu Asp Gln Lys Ala Arg Glu Ala Phe Arg Arg
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 19:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Ser Val Glu Leu Arg Glu Leu Tyr Val Met Glu Ile Ser Asp Asn Pro
1 5 10 15
Gly Val His Glu Ala Gly Glu Pro Glu Phe Lys
20 25

(2) INFORMATION FOR SEQ ID NO: 20:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

RAW SEQUENCE LISTING DATE: 11/21/2001
PATENT APPLICATION: US/09/818,066 TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt
Output Set: N:\CRF3\11212001\I818066.raw

```
294 Leu Ile Asp Arg Thr Arg Ile Val Ile Val Pro Ser Leu Asn Pro Asp
295  1               5                  10                  15
297 Gly Arg Ile Ala
298              20
300 (2) INFORMATION FOR SEQ ID NO: 21:
302      (i) SEQUENCE CHARACTERISTICS:
303           (A) LENGTH: 22 amino acids
304           (B) TYPE: amino acid
305           (D) TOPOLOGY: linear
307     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:
309 Ser Leu Leu Ser His Glu Phe Gln Asp Glu Thr Asp Thr Glu Glu Glu
310  1               5                  10                  15
312 Thr Leu Tyr Ser Ala Lys
313              20
315 (2) INFORMATION FOR SEQ ID NO: 22:
317      (i) SEQUENCE CHARACTERISTICS:
318           (A) LENGTH: 13 amino acids
319           (B) TYPE: amino acid
320           (D) TOPOLOGY: linear
322     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:
324 Val Glu Glu Gly Lys Val Pro Val Leu Asn Thr Pro Asp
325  1               5                  10
327 (2) INFORMATION FOR SEQ ID NO: 23:
329      (i) SEQUENCE CHARACTERISTICS:
330           (A) LENGTH: 22 amino acids
331           (B) TYPE: amino acid
332           (D) TOPOLOGY: linear
334     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:
336 Glu Leu Tyr Val Met Glu Ile Ser Asp Asn Pro Gly Val His Glu Ala
337  1               5                  10                  15
339 Gly Glu Pro Glu Phe Lys
340              20
342 (2) INFORMATION FOR SEQ ID NO: 24:
344      (i) SEQUENCE CHARACTERISTICS:
345           (A) LENGTH: 66 base pairs
346           (B) TYPE: nucleic acid
347           (C) STRANDEDNESS: single
348           (D) TOPOLOGY: linear
350     (ix) FEATURE:
355           (A) NAME/KEY: misc_feature
352           (B) LOCATION: 3..3
353           (D) OTHER INFORMATION: /note= "R is A or G."
375     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:
377 GARCTNTAYG TNATGGARAT WAGYGAYAAY CCNGGNGTNC AYGARGCNGG NGARCCNGAR      60
379 TTYAAR                                                               66
381 (2) INFORMATION FOR SEQ ID NO: 25:
383      (i) SEQUENCE CHARACTERISTICS:
384           (A) LENGTH: 66 base pairs
385           (B) TYPE: nucleic acid
```

VERIFICATION SUMMARY DATE: 11/21/2001
PATENT APPLICATION: US/09/818,066 TIME: 12:09:11

Input Set : N:\Crf3\RULE60\09818066.txt
Output Set: N:\CRF3\11212001\I818066.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:1982 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:1986 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:1990 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:1994 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:1998 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2002 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2006 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2010 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2014 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2018 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2022 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2026 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2030 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2034 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2038 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2042 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2046 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2050 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2054 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2058 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2062 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2066 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2070 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2074 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2175 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2179 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2183 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2187 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2191 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2195 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2199 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2203 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2207 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2211 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2369 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=74